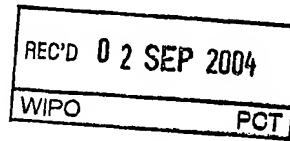


10/510391

Rec'd PCT/PTO 07 OCT 2004

PCT/AU2004/001087



Patent Office  
Canberra

I, LEANNE MYNOTT, MANAGER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. 2003904350 for a patent by SILVERBROOK RESEARCH PTY LTD. as filed on 15 August 2003.

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

WITNESS my hand this  
Twenty-fifth day of August 2004

LEANNE MYNOTT  
MANAGER EXAMINATION SUPPORT  
AND SALES



# SYSTEMS, METHODS AND APPARATUS

Jon Napper and Paul Lapstun  
 [{jon.napper,paul.lapstun}@silverbrookresearch.com](mailto:{jon.napper,paul.lapstun}@silverbrookresearch.com)

Silverbrook Research Pty Ltd  
393 Darling Street, Balmain NSW Australia  
30 June 2003

## 1 Introduction

The increasing use of pen computing and the emergence of paper-based interfaces to networked computing resources [10,11] has highlighted the need for techniques to search raw digital ink. Pen-based computing allows users to store data in the form of digital ink notes and annotations, and subsequently search this data using hand-written or hand-drawn queries. However, searching raw digital ink is more difficult than traditional text searching due to variations and inconsistencies in the production of handwriting and hand-drawn images, and thus methods for improving search accuracy using domain-specific knowledge, constraints, and contextual information are valuable. This document discusses a number of novel techniques for improving the accuracy of digital ink searching.

### 1.1 Cross-References

Various methods, systems and apparatus relating to the present invention are disclosed in the following co-pending applications filed by the applicant or assignee of the present invention. The disclosures of all of these co-pending applications are incorporated herein by cross-reference.

5 October 2002:

Australian Provisional Application 2002952259, "Methods and Apparatus (NPT019)".

15 October 2002:

PCT/AU02/01391, PCT/AU02/01392, PCT/AU02/01393, PCT/AU02/01394 and PCT/AU02/01395.

26 November 2001:

PCT/AU01/01527, PCT/AU01/01528, PCT/AU01/01529, PCT/AU01/01530 and PCT/AU01/01531.

11 October 2001:

PCT/AU01/01274.

14 August 2001:

PCT/AU01/00996.

27 November 2000:

PCT/AU00/01442, PCT/AU00/01444, PCT/AU00/01446, PCT/AU00/01445, PCT/AU00/01450, PCT/AU00/01453, PCT/AU00/01448, PCT/AU00/01447, PCT/AU00/01459, PCT/AU00/01451, PCT/AU00/01454, PCT/AU00/01452, PCT/AU00/01443, PCT/AU00/01455, PCT/AU00/01456, PCT/AU00/01457, PCT/AU00/01458 and PCT/AU00/01449.

20 October 2000:

PCT/AU00/01273, PCT/AU00/01279, PCT/AU00/01288, PCT/AU00/01282, PCT/AU00/01276, PCT/AU00/01280, PCT/AU00/01274, PCT/AU00/01289, PCT/AU00/01275, PCT/AU00/01277, PCT/AU00/01286, PCT/AU00/01281, PCT/AU00/01278, PCT/AU00/01287, PCT/AU00/01285, PCT/AU00/01284 and PCT/AU00/01283.

15 September 2000:

PCT/AU00/01108, PCT/AU00/01110 and PCT/AU00/01111.

30 June 2000:

PCT/AU00/00762, PCT/AU00/00763, PCT/AU00/00761, PCT/AU00/00760, PCT/AU00/00759, PCT/AU00/00758, PCT/AU00/00764, PCT/AU00/00765, PCT/AU00/00766, PCT/AU00/00767, PCT/AU00/00768, PCT/AU00/00773, PCT/AU00/00774, PCT/AU00/00775, PCT/AU00/00776, PCT/AU00/00777, PCT/AU00/00770, PCT/AU00/00769, PCT/AU00/00771, PCT/AU00/00772, PCT/AU00/00754, PCT/AU00/00755, PCT/AU00/00756 and PCT/AU00/00757.

24 May 2000:

PCT/AU00/00518, PCT/AU00/00519, PCT/AU00/00520, PCT/AU00/00521, PCT/AU00/00522, PCT/AU00/00523, PCT/AU00/00524, PCT/AU00/00525, PCT/AU00/00526, PCT/AU00/00527, PCT/AU00/00528, PCT/AU00/00529, PCT/AU00/00530, PCT/AU00/00531, PCT/AU00/00532, PCT/AU00/00533, PCT/AU00/00534, PCT/AU00/00535, PCT/AU00/00536, PCT/AU00/00537, PCT/AU00/00538, PCT/AU00/00539, PCT/AU00/00540, PCT/AU00/00541, PCT/AU00/00542, PCT/AU00/00543, PCT/AU00/00544, PCT/AU00/00545, PCT/AU00/00547, PCT/AU00/00546, PCT/AU00/00554, PCT/AU00/00556, PCT/AU00/00557, PCT/AU00/00558, PCT/AU00/00559, PCT/AU00/00560, PCT/AU00/00561, PCT/AU00/00562, PCT/AU00/00563, PCT/AU00/00564, PCT/AU00/00565, PCT/AU00/00566, PCT/AU00/00567, PCT/AU00/00568, PCT/AU00/00569, PCT/AU00/00570, PCT/AU00/00571, PCT/AU00/00572, PCT/AU00/00573, PCT/AU00/00574, PCT/AU00/00575, PCT/AU00/00576, PCT/AU00/00577, PCT/AU00/00578, PCT/AU00/00579, PCT/AU00/00581, PCT/AU00/00580, PCT/AU00/00582, PCT/AU00/00587, PCT/AU00/00588, PCT/AU00/00589, PCT/AU00/00583, PCT/AU00/00593, PCT/AU00/00590, PCT/AU00/00591, PCT/AU00/00592, PCT/AU00/00594, PCT/AU00/00595, PCT/AU00/00596, PCT/AU00/00597, PCT/AU00/00598, PCT/AU00/00516, PCT/AU00/00517 and PCT/AU00/00511.

## 1.2 Digital Ink Definition

Digital ink is a digital representation of the information generated by a pen-based input device. Generally, digital ink is structured as a sequence of strokes that begin when the pen device makes contact with a drawing surface and ends when the pen device is lifted. Each stroke comprises a set of sampled coordinates that define the movement of the pen whilst the pen is in contact with the drawing surface.

## 1.3 Digital Ink Searching

The traditional method of searching handwritten data is to first convert the ink database and corresponding query to text using pattern recognition techniques, and then to match the query text with the text in the database. Fuzzy text searching methods have been described [13] that perform text matching in the presence of character errors similar to those produced by handwriting recognition systems.

However, handwriting recognition accuracy remains low, and the number of errors introduced by recognition (both for the database entries and for the handwritten query) means that this technique does not work well. The process of converting ink into text results in the loss of a significant amount of information regarding the general shape and dynamic properties of the ink. For example, some letters (e.g. 'u' and 'v', 'v' and 'r', 'f' and 't', etc.) are handwritten with a great deal of similarity in shape. Additionally, in many handwriting styles (particularly cursive writing), the identification of individual characters is highly ambiguous.

Digital ink searching refers to the process of searching through a continuous stream of digital ink for patterns that most closely match the input query according to some similarity metric. Direct matching on raw digital ink allows shape information to be considered during the search procedure, and does not require character or word segmentation to be performed. Techniques for digital ink searching have been proposed in [11,12,13,14,15,16,17,18,19,20,21].

#### **1.4 Digital Ink Search Applications**

A number of highly desirable applications are made possible by the combination of digital ink persistence and digital ink searching, including the ability to search annotations, notes, comments, and other handwritten information for keywords or phrases. The digital ink searching procedure is not limited to simply matching the query text, as additional attributes can be used to more accurately specify the desired information. Examples of these attributes include: date and time of writing, the identity of pen used to produce the writing, geographic location where the writing took place, application with which the writing is associated (e.g. electronic mail or notebook), type of field that contains the writing (e.g. a text input field, a drawing field), the location of the annotation or text on the page, and so on.

Pen-based queries also allow searching for information other than handwriting. Hand-drawn picture searching can be used to locate drawings and diagrams in a notebook, and can be used to search a collection of digital images. As an example, a hand-drawn picture query could be used to search an online photo album or commercial image library for pictures that contains a desired visual feature or set of visual features [15].

#### **1.5 Detailed Description of the Preferred Embodiments**

In the preferred embodiment, the invention is configured to work with the Netpage networked computer system, a detailed description of which is given in our co-pending applications, including in particular PCT application WO0242989 entitled "Sensing Device" filed 30 May 2002, PCT application WO0242894 entitled "Interactive Printer" filed 30 May 2002, PCT application WO0214075 "Interface Surface Printer Using Invisible Ink" filed 21 February 2002, PCT application WO0242950 "Apparatus For Interaction With A Network Computer System" filed 30 May 2002, and PCT application WO03034276 entitled "Digital Ink Database Searching Using Handwriting Feature Synthesis" filed 24 April 2003. It will be appreciated that not every implementation will necessarily embody all or even most of the specific details and extensions described in these applications in relation to the basic system. However, the system is described in its most complete form to assist in understanding the context in which the preferred embodiments and aspects of the present invention operate.

In brief summary, the preferred form of the Netpage system provides an interactive paper-based interface to online information by utilizing pages of invisibly coded paper and an optically imaging pen. Each page generated by the Netpage system is uniquely identified and stored on a network server, and all user interaction with the paper using the Netpage pen is captured, interpreted, and stored. Digital printing technology facilitates the on-demand printing of Netpage documents, allowing interactive applications to be developed. The Netpage printer, pen, and network infrastructure provide a paper-based alternative to traditional screen-based applications and online publishing services, and supports user-interface functionality such as hypertext navigation and form input.

Typically, a printer receives a document from a publisher or application provider via a broadband connection, which is printed with an invisible pattern of infrared tags that each encodes the location of the tag on the page and a unique page identifier. As a user writes on the page, the imaging pen decodes these tags and converts the motion of the pen into digital ink. The digital ink is transmitted over a wireless channel to a relay base station, and then sent to the network for processing and storage. The system uses a stored description of the page to interpret the digital ink, and performs the requested actions by interacting with an application.

Applications provide content to the user by publishing documents, and process the digital ink interactions submitted by the user. Typically, an application generates one or more interactive pages in response to user input, which are transmitted to the network to be stored, rendered, and finally printed as output to the user. The Netpage system allows sophisticated applications to be developed by providing services for document publishing, rendering, and delivery, authenticated transactions and secure payments, handwriting recognition and digital ink searching, and user validation using biometric techniques such as signature verification.

## 2 Domain-Specific Specialization

Many of the digital ink searching algorithms described above are designed to search a specific type of digital ink. For example, the systems proposed in [23,24] are most effective when searching printed or cursive handwritten Latin-script text, whilst [12,14,16] describe techniques for searching hand-drawn pictures. Similarly, systems can be developed that are optimised for searching other specific types of digital ink, such as oriental handwritten characters, technical drawings, or hand-drawn equations.

In most cases, systems designed to search a specific form of digital ink will achieve greater accuracy than general-purpose digital ink searching methods, since these systems are able to utilize domain-specific knowledge when designing the ink searching algorithms. Knowledge of the expected digital ink format will influence the selection of segmentation techniques, pre-processing and normalization, the pattern primitives used (e.g. stroke, sub-stroke, stroke group, bitmap image, etc.), the extracted feature set, the matching algorithm, the similarity metric, and so on. The steps required to perform digital ink searching using specialization are given in Figure 1.

### 2.1 Specialization Examples

For searching cursive Latin-script handwriting, techniques can be developed to exploit the key characteristics of this type of writing, such as the powerful discriminatory influence of ascender and descender elements (e.g. "bdfghjklpqty"), the existence of specific zones within the writing (base lines and core lines), and the relatively stable ordering of the handwritten strokes (at least within the writing of a single author). Additional high-level information can also be utilized, such as the expectation that the writing will be clustered into approximately linear lines that contain groups of strokes representing words and letters. Further specialization is possible if it is known that the matching digital ink is largely numeric (e.g. a phone number), since digits are usually drawn consistently, being well segmented (no ligatures) and with a regularity of character height. Specialized search strategies are also possible for handwritten text that contains only upper-case letters.

However, the requirements for accurately searching hand-drawn pictures and scribbles are significantly different, and most of the key discriminatory characteristics of handwriting are not available. Hand-drawn picture search algorithms must be stroke order and stroke direction insensitive, due to the large number of different ways the same picture may be drawn. Generally, the algorithm must also be rotationally insensitive, since drawings can be made at arbitrary orientations on a page. To improve accuracy, picture searching algorithms may exploit the fact that most drawings are rendered using an aggregation of line and shape primitives that may be used to decompose the image into a canonical form useful for similarity matching.

Other domain-specific specializations for digital ink search can also be made. For example, systems for searching oriental handwritten characters can utilize the highly accurate character segmentation techniques that have been developed for oriental character recognition systems [26]. In addition to this, they may exploit the fact that the characters are generally composed from a small set of primitive radicals, whilst compensating for the potentially large stroke-order variation that can occur during writing.

Additional specializations exist for other types of digital ink data, such as hand-drawn equations, diagrams, and charts. In general, specializations can be made for any type of digital ink data that contains a structure or regularity that may be exploited to provide improved discriminatory features. An awareness of the constraints and expected deviation of the data can be used to differentiate noise from information, and thus provide a more accurate similarity metric.

### 2.2 Using Specialized Searching

Having a set of specialized searching strategies is only useful if it can be accurately determined when each particular strategy should be used. In the simplest case, the determination is made at a system level; for example, allowing a system administrator to select Latin-script based searching or oriental character searching depending on the location or expected users of the system. It is also possible for this decision to be made automatically, given the existence of metrics that can accurately differentiate between Latin-based and oriental scripts [25,27]. Similar techniques exist to differentiate written text from hand-drawing images.

A more flexible system allows individual segments of digital ink to be labelled as a specific digital ink type, and subsequently searched using algorithms specialized for that particular type. For example, the system may allow a user to indicate that they generally write using a specific language (e.g. in English or Chinese) or writing style (e.g. cursive, printed, upper-case, or mixed) and this information can be used to select the appropriate ink searching mechanism. In addition to this, the system may allow the user to manually indicate the type of digital ink being generated. For example, the user could use a number of different pens (e.g. one for handwriting text and another for drawing pictures) allowing the system to discriminate between the different ink types. Alternatively, gestures or other user-initiated actions could be used to label ink data.

Another approach to specialized digital ink searching is to require the manual selection of the search method when the search query is generated. For example, if the user wishes to search for English handwritten text, they write their text query, and then indicate to the system that an English handwritten text search should be performed using the specified query. Similarly, if the user wishes to search for a hand-drawn picture, they draw their query and indicate to the system to perform a drawing search. Since most digital ink searching systems perform some kind of pre-processing or indexing at the time of ink generation (rather than when the query is generated) to ensure a fast response to ink search queries, delaying the search strategy decision until the point at which the search is initiated means that either:

- the ink data must be pre-processed multiple times and stored in multiple formats (i.e. once for each search strategy), or
- the pre-processing must be delayed until the search is initiated (thus increasing the time it takes to generate the search results).

The improvement in the accuracy of the ink search may justify the increased resource utilization required by this technique.

### 3 Specialization Using Context Information

In addition to the techniques described above, the application of specialized digital ink searching techniques can be determined from the context (i.e. the contents of the page or document on which the ink was written) of the digital ink. Interpreting the information contained in the layout and definition of a document can guide the selection of the ink search strategy.

#### 3.1 Language/Script Identification

It is reasonable to assume that annotations and comments made on a printed document will usually be written in the same language as the text contained in the document itself. Thus, if the natural language of a document (i.e. the language that the text in the document was written in) can be determined, specialized ink search strategies can be used to search digital ink annotations contained on that document.

Many document formats allow the explicit definition of the natural language of the document. For example, in HTML/XHTML [6,7] the "lang" attribute can be used:

```
<HTML lang="en" dir="rtl"></HTML>
```

where the language is identified by a two-letter code (e.g. "en" for English, "es" for Spanish, etc.) as specified in [1,2]. The example also shows the ability to specify the text direction ("dir") as right-to-left ("rtl") or left-to-right ("ltr"), another assumed characteristic of the digital ink that can be used when performing digital ink searching. Similarly, in the XML/XFORMS document specification "a special attribute named "xml:lang" may be inserted in documents to specify the language used in the contents and attribute values of any element in an XML document" [4]:

```
<TITLE xml:lang="fr">XForms en XHTML</TITLE>
```

The Adobe Portable Document Format (PDF) defines the "Lang" attribute, a "language identifier specifying the natural language for all text" [5]. The identifier can be used in the document catalog (thus specifying the language of the entire document), in any structured element, or in marked-content sequences:

```
/Span << /Lang (fr) >>
  BDC
    (Bonjour.) Tj
  EMC
```

Documents may also use the Dublin Core metadata element set, "a standard for cross-domain information resource description" [28] that identifies the language associated with a resource using the standard language codes [2,3]. Dublin Core metadata conforms to the World Wide Web Consortium (W3C) Resource Description Framework, and can be used with HTML and XML documents [29,30].

If a document format does not allow the specification of the document language, or the language specification attribute is missing, the language of the document may be inferred using other techniques. For example, the use of a particular font will often imply that the document was authored in a particular language or script. In some document formats (such as PDF [5]), font objects contain a language attribute that indicates the natural language of the font. In addition to this there exist techniques that allow the language of a document to be accurately determined using dictionaries.

Note that some specialized digital ink searching techniques are optimised for a specific script (e.g. Latin characters, Oriental characters, Arabic characters, etc.) that includes a group of languages, rather than being language specific. Obviously, any technique that exploits language identification for specialization can also be used for language script based specialization, since identification of the language script is usually trivial once the language is known.

### 3.2 Field Labels

Documents and forms that require data to be entered, either using a keyboard for screen-based applications or handwritten for pen computing or paper documents, must give the user some indication of the type of information that is required. This is usually done by labelling each data input area (or field) with a descriptive identifier, for example, "First Name", "Last Name", "Address", "Phone Number", and so on. For printed forms, this information appears as printed text on the paper, while online (i.e. computer-based) documents usually contain this information as a visible text entry defined in the structured description of the form.

The information contained in the field labels described above can be used to determine the digital ink searching strategy to use for the digital ink contained in the field. This is done by first associating each field label with the appropriate data entry region by analysing the form description to associate labels with data entry regions. Once each label is associated with an entry field, a table of previously defined field label strings is searched (possibly using regular expression matching) and the corresponding ink type and associated ink search strategy is found. The following are some example ink types and associated field titles:

Ink Type	Field Label
Text	First Name, Given Name, Surname, Family Name, Address, Suburb, Town, State, Country, Region, Email Address
Numeric	Phone number, Age, Number, Size, Count, Zip Code, Post Code, Date, Time, Credit Card Number, Customer Number
Drawing	Picture, Drawing, Image, Diagram

### **3.3 Field Attributes**

In addition to the field type, form definitions often contain information regarding the type of data that should be entered in each field. This information is usually contained in attributes that are associated with a specific field. For example, some input field types have a flag indicating that the value entered must be numeric. A digital ink searching system can use this information to select a numeric search strategy for ink contained in the associated data input area.

In addition to using standard form field attributes to improve the accuracy of digital ink searching, digital ink search specific information can be added to fields using custom attributes. This information is only used if the document is processed using a digital ink searching system; the document can still be used normally where required (e.g. printed or displayed in web browser) since processing systems generally ignore the unrecognised custom attributes. However, if digital ink searching is required, the custom parameters can be used to improve the accuracy of the search results.

## **4 Conclusion**

A number of techniques to improve the accuracy of digital ink search were discussed, including a method of selecting a digital ink searching strategy from a set of specialized strategies based on the expected ink data type. Examples of specialized digital ink searching strategies were given, along with a number of methods for integrating these strategies into a system. In addition to this, methods for selecting digital ink search strategies based on the layout and definition of a document was discussed.

## 6 References

- [1] H. Alvestrand, "Tags for the Identification of Languages", RFC 3066, January 2001.
- [2] International Organization for Standardization, "Codes for the Representation of Names of Languages", ISO 639, 1988.
- [3] International Organization for Standardization, "Codes for the Representation of Names of Countries and Their Subdivisions", ISO 3166.
- [4] World Wide Web Consortium, "XForms 1.0", W3C Working Draft 21, August 2002.
- [5] Adobe Systems Incorporated, "Acrobat Core API Reference", Version Acrobat 5.0, December 2001.
- [6] World Wide Web Consortium (W3C), "HTML 4.01 Specification", W3C Recommendation, 24 December 1999.
- [7] World Wide Web Consortium (W3C), "XHTML 1.0 The Extensible HyperText Markup Language (Second Edition)", W3C Recommendation, 1 August 2002.
- [8] International Organization for Standardization (ISO), "Information processing - Text and office systems - Standard Generalized Markup Language (SGML)", ISO 8879, 1986.
- [9] P. Lapstun, Netpage System Overview, Internal Report, Silverbrook Research Pty Ltd, 6th June, 2000.
- [10] Anoto, "Anoto, Ericsson, and Time Manager Take Pen and Paper into the Digital Age with the Anoto Technology", Press Release, 6th April, 2000.
- [11] Y. Chans, Z. Lei, D. Lopresti, and S. Kung, "A Feature Based Approach For Image Retrieval by Sketch", Proceedings of SPIE Volume 3229: Multimedia Storage and Archiving Systems II, 1997
- [12] D. Lopresti and A. Tomkins, "Temporal-Domain Matching of Hand-Drawn Pictorial Queries", Handwriting and Drawing Research: Basic and Applied Issues, IOS Press, pp. 387-401, 1996.
- [13] D. Lopresti and A. Tomkins, "Block Edit Models for Approximate String Matching", Proceedings of the 2nd Annual South American Workshop on String Processing, pp. 11-26.
- [14] D. Lopresti, A. Tomkins, and J. Zhou, "Algorithms for Matching Hand-Drawn Sketches", Proceedings of the 5th International Workshop on Frontiers in Handwriting Recognition, pp. 223-238, 1995.
- [15] A. Del Bimbo, P. Pala, and S. Santini, "Image Retrieval by Elastic Matching of Shapes and Image Patterns", Proceedings of IEEE Multimedia, pp. 215-218, 1996.
- [16] D. Lopresti and A. Tomkins, "Approximate Matching of Hand-Drawn Pictograms", 3rd International Workshop on Frontiers in Handwriting Recognition, 1993.
- [17] I. Pavlidis, R. Singh, and N. Papanikopoulos, "Recognition of On-Line Handwritten Patterns Through Shape Metamorphosis", Proceedings of the 13th International Conference on Pattern Recognition, Vol. 3, pp 18-22, 1996.
- [18] L. Schomaker, L. Vuurpijl, and E. de Leau, "New Use for the Pen: Outline-Based Image Queries", Proceedings of the 5th International Conference on Document Analysis and Recognition, pp. 293-296, 1999.
- [19] S. Muller, S. Eickeler, and G. Rigoll, "Multimedia Database Retrieval Using Hand-Drawn Sketches", 5th International Conference on Document Analysis and Recognition, Bangalore, India, September 1999.
- [20] R. Manmatha, C. Han, E. Riseman, and W. Croft, "Indexing Handwriting Using Word Matching", Proceedings of the First ACM International Conference on Digital Libraries, pp. 151-159, 1996.
- [21] A. Poon, K. Weber, and T. Cass, "Scribbler: A Tool for Searching Digital Ink", Proceedings of the ACM Computer-Human Interaction, pp.58-64, 1994.
- [22] D. Lopresti and A. Tomkins, "Pictographic Naming", Proceedings of the INTERCHI '93 Conference, 1993.

- [23] I. Kamel, "Fast Retrieval of Cursive Handwriting", Proceedings of the 5th International Conference on Information and Knowledge Management, Rockville, MD USA, November 12-16, 1996.
- [24] J. Napper, P. Lapstun, and K. Silverbrook, "Digital Ink Searching", Internal Report, Silverbrook Research Pty Ltd, December 2001.
- [25] U. Pal, and B. Chaudhuri, "Automatic Identification of English, Chinese, Arabic, Devnagari and Bangala Script Line", Sixth International Conference on Document Analysis and Recognition, September 2001.
- [26] C. Hong, G. Loudon, Y. Wu, R. Zitserman, "Segmentation and Recognition of Continuous Handwriting Chinese Text", Advances in Oriental Document Analysis and Recognition Techniques, World Scientific Publishing, pp. 223-232, 1998.
- [27] L. Lam, J. Ding, C. Suen, "Differentiating Between Oriental and European Scripts by Statistical Features", Advances in Oriental Document Analysis and Recognition Techniques, World Scientific Publishing, pp. 63-80, 1998.
- [28] Dublin Core Metadata Initiative, "Dublin Core Metadata Element Set, Version 1.1: Reference Description", June 2003.
- [29] J. Kunze, "Encoding Dublin Core Metadata in HTML", RFC 2731, December 1999.
- [30] D. Beckett, E. Miller, D. Brickley, "Expressing Simple Dublin Core in RDF/XML", Dublin Core Metadata Initiative, July 2002.

## 5 Figures

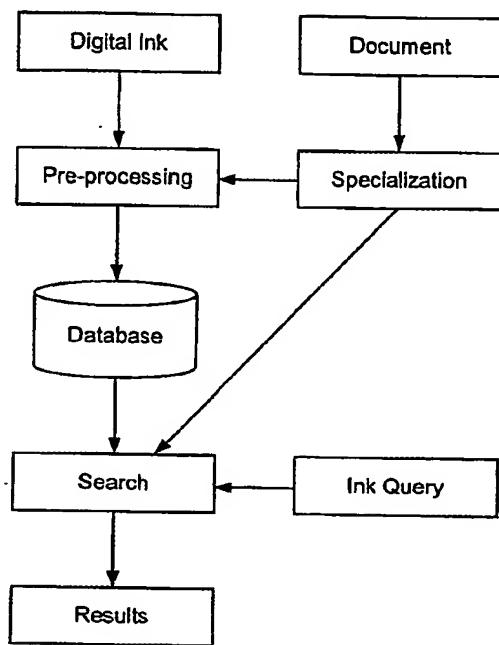


Figure 1. Digital Ink Searching Using Specialization